DOCUMENT RESUME

ED 205 560 TM 810 442

AUTHOR Linn, Marcia C.; And Others
TITLE Is It Formal If It's Not Physics? (The Influence of Laboratory and Naturalistic Content on Formal Reasoning).
INSTITUTION California Univ., Berkeley. Lawrence Hall of Science.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE [80]
GRANT SED77-18914
NOTE 35p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Processes; *Cognitive Style; *Critical Thinking; Physics; *Problem Solving; *Program Content; *Scientific Principles; Secondary Education
IDENTIFIERS *Formal Operations; Laboratory Content; Naturalistic Content; *Piagetian Tasks; Variables

ABSTRACT

The study reported here extends the investigation begun by Piaget of content effects to laboratory and naturalistic tasks. If the accuracy and completeness of the subjects' expected variables is related to content, then performance on laboratory and naturalistic tasks should differ systematically. The subjects' expected variables were measured by having subjects name variables and by a group survey. Use of the controlling variables strategy was measured by four questions about constructing, criticizing, planning, and analyzing experimental results. Subjects of different ages and socioeconomic groups were used. Results revealed a strong influence of content on reasoning performance. Contrary to Piaget's writings, content effects were systematic. In addition, by choosing physics problems and questions which deemphasized content, Piaget avoided many content influences. Thus, it is suggested that educational implications generated from Piagetian theory may be incomplete: content effects require more emphasis. (Author/GK)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED205560

# Is It Formal If It's Not Physics?

## (The Influence of Laboratory and Naturalistic Content on Formal Reasoning)

Marcia C. Linn

Cathy Clement

Steven Pulos

Lawrence Hall of Science
University of California
Berkeley, CA 94720

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. Linn

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

This material is based upon research supported by the National Science Foundation under grant No. SED77-18914 directed by Marcia C. Linn. Any opinions, findings, and conclusions or recommendations expressed in this document are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The authors would like to thank Darlene Clement, Lee Cronbach, Tina deBenedictis, Kevin Delucchi, Elizabeth Stage, Cary Sneider, and Pat Sullivan for critical comments and suggestions. Help with test development and administration by Harbor Fraser, Lois Foley, and Janet Rocha is appreciated.

TM 810 442

## Abstract

Formal reasoning has been studied extensively with tasks that have physics or laboratory content—this research investigates the effect of content on formal reasoning. Ninety, thirteen, fifteen, and seventeen-year-olds received both laboratory and naturalistic content tasks which required ability to control variables. Expectations about the variables in each task were measured. Results revealed that 8% to 15% of the variance in performance was associated with task content. Content effects were shown to reflect expectations about task variables.

## Is It Formal If It's Not Physics?

### (The Influence of Laboratory and Naturalistic Content on Formal Reasoning)

Formal reasoning has been studied predominately with tasks that have physics (or laboratory) content, raising the question: Is it formal if it's not physics? Inhelder and Piaget's (1958) monumental work on formal reasoning focused on the strategies required for formal reasoning, such as the strategy for controlling variables, not on the task content. In this paper, we examine how laboratory and naturalistic content influence the application of formal reasoning strategies.

Inhelder and Piaget (1958) hypothesized that content would unsystematically influence performance; their research demonstrated little concern for content. First, their tasks came predominately from physics -- they did not compare tasks from different content domains. Second, they confounded strategy with content, for example, using one content (bending rods) to measure the controlling variables strategy and another (balance beam) to measure the proportional reasoning strategy.

Piaget has written very little about content, indicating his limited concern. In 1971, he discussed "resistances" due to content which influence performance. He said that effects of resistances were easy to explain after they had occured but hard to predict. Piaget believed that content idiosyncratically influences performance. One idiosyncratic influence described by Piaget is expertise: To explain why formal reasoning was less prevalent than he hypothesized, Piaget (1972) cited the effect of expert content knowledge. Piaget said individuals were more likely to reason formally in their area of expertise than in other areas. For example, an auto mechanic might display formal reasoning when diagnosing a defect in an auto but not when diagnosing a child's allergy.

In contrast to Piaget, who did not research content and thought content effects were idiosyncratic, we hypothesized that content systematically influences performance. For any task, the subject expects that certain variables need to be considered; we call these the "subjects' expected variables". The variables that the experimenter selects for a task (the "experimenters' comprehensive variables") may differ from the subjects' expected variables. We hypothesized that content will influence formal reasoning task performance because, for a given content, subjects have particular expectations regarding the variables.

Systematic differences between the subject's expected variables and the experimenter's comprehensive variables can influence performance in the following ways: Subjects may consider fewer variables than the experimenter. Expectations may also be inaccurate; for example, most people inaccurately expect the weight of the bob to influence the oscillation of the pendulum because weight influences other phenomena.

While Piaget only studied tasks with laboratory or physics content, such as Bending Rods, we investigated content effects by comparing a variety of laboratory and naturalistic tasks. Naturalistic tasks have content from frequently encountered situations such as determining which is the best toothpaste to buy or how to get the best mileage. In the naturalistic tasks as in the laboratory tasks, we asked subjects about the design of controlled experiments.

We hypothesized that the completeness and accuracy of the subjects' expected variables would be influenced by content. Specifically, we anticipated that the subjects' expected variables for laboratory tasks would differ from the subjects' expected variables for naturalistic tasks. We planned to demonstrate systematic content effects by relating content differences in the subjects' expected variables to content differences in reasoning.

Laboratory and naturalistic tasks assessed a single formal reasoning strategy: the controlling variables strategy. (Using the controlling variables strategy means designing experiments where one variable at a time is manipulated while others are kept the same.) Subjects' use of controlling variables strategy was assessed by questions which involved constructing, criticizing, planning, and analyzing experiments. Previous research (e.g., Linn, 1978) has shown that these questions reveal different aspects of strategy knowledge. In this study, questions about the controlling variables strategy were selected to reveal the influence of the subjects' expected variables on strategy use.

Preliminary Studies of Content Effects

Two preliminary studies of content effects in formal reasoning lay the groundwork for the hypotheses investigated in the present paper. To investigate Piaget's (1972) statements that people are more likely to reason formally in their area of expertise, Pulos and Linn (1978) studied two groups of seventh graders who had different levels of expertise for controlling variables tasks with fishing and laboratory science content. One group lived on a river delta and often went fishing. The other group was enrolled in an experiential science program. The authors hypothesized that the subjects' expected variables would be most accurate and most complete in their area of expertise and would lead to superior performance by experts. They found that subjects performed better on controlling variables tasks in their area of expertise and performed equivalently on a neutral content task where neither group was expert. Preliminary evidence for a relationship between subjects' expected variables and subjects' performance was found. These results suggest that expertise influences the ability to control variables.

Linn and Swiney (Note 1), in a second study of content effects with eleventh graders, investigated the relationship between the subject's expected variables, and the variables the subject actually utilized on a laboratory controlling

task. They measured the subject's expected variables by having subjects examine the task apparatus and name the variables they expected would influence the outcome. Subjects were presumed to name their expected variables. Subjects were then told the experimenter's comprehensive variables. Subjects demonstrated that they utilized a variable by correctly controlling or investigating it in an experiment. Comparison of the subjects' expected variables and the subjects' utilized variables revealed that subjects utilized their expected variables significantly more often than other variables. Subjects appeared to answer the controlling variables questions as if only their expected variables were important -- subjects ignored the experimenter's hypothetical variables.

The Present Study

The study reported in this paper extends the investigation of content effects to laboratory and naturalistic tasks. If the accuracy and completeness of the subjects' expected variables is related to content, then performance on laboratory and naturalistic tasks should differ systematically. The subjects expected variables were measured by having subjects name variables and by a group survey called What is Your Opinion? Use of the controlling variables strategy was measured by four questions about constructing, criticizing, planning, and analyzing experimental results. In contrast to earlier studies which used a single age and socioeconomic group, this study used three ages (13, 15, and 17) and three schools differing in socioeconomic status.

## Methods

### Subjects

Ninety 13, 15, and 17 year olds in three different school districts (ten per age per district) participated. Four subjects moved during the study leaving a final sample of 86. The districts represented three socioeconomic status (SES) groups: upper middle class, middle class, and lower middle class. The upper middle class district was suburban: families owned their own homes and many students drove their own cars to school. The middle class district was urban: families lived in single family dwellings, apartments, and condominiums. The lower middle class district was semirural: families lived in small single family dwellings, apartments, and condominiums, and many adults were employed in adjacent factories. Quality of the science programs in the three districts correlated directly with SES.

The What is Your Opinion? survey of task expectations was administered to a larger sample (N = 900) of which the subjects in this study were a randomly selected subsample.

### Controlling Variables Tasks

Six naturalistic and four laboratory tasks were devised. The laboratory tasks had apparatus but the naturalistic tasks did not. Formats to communicate naturalistic tasks were pilot tested and three were selected. Some or all of the five questions (constructing, criticizing, planning, analyzing, naming) were asked for each task. Scoring was pass/fail, although for some questions, additional record was made of specific responses. Table 1 summarizes the format used, type of questions asked, and number of questions asked for each task. Tasks were administered in two forty-minute interviews; in the first interview, the order of the tasks was systematically varied across subjects; in the second interview, order was constant.

Insert Table 1 about here

## Tasks

The criteria and variables for each of the ten tasks were as follows:

Laboratory tasks:

1. Bending Rods, a task adapted from Inhelder and Piaget (1958) and described by Linn and Swiney (Note 1), allows investigation of the bendability of rods. The five variables are: the material of the rod, the thickness of the rod, the shape of the rod, the length of the rod, and the size of the weight hung from the rod.

2. Springs, a task described by Linn and Rice (1979), is parallel to Bending Rods, but subjects investigate the expansion of springs. Variables are: the material of the spring, the cross-section of the spring, the length of the spring, the thickness of the wire, and the weight hung from the spring.

3. Spinning Wheels, a task adapted by Case (1974) from Inhelder and Piaget (1958), allows investigation of centrifugal force. Spheres are placed on a wheel which can be rotated; the question to be answered is which sphere will fly off the wheel first. The variables are: the size of the sphere, the material of the sphere, the size of the hole in which the sphere rests, and the distance of the hole from the center of the wheel.

4. Runway, allows investigation of linear forces on a toy sled traveling down a ramp. Factors influencing the distance travelled by the sled which is pushed by a marble released from the top of the ramp are investigated. Variables include: the type of passengers, the number of passengers, the type of sled, the size of the marble used to push the sled, the weight of the marble, and the height of release of the marble.

Naturalistic tasks:

5. In Toothpaste, subjects investigate what sells toothpaste. The variables are: the presence or absence of flouride, the flavor, the presence or absence of a dentist's recommendation, the cost, and the amount of advertising.

6. In Autos, subjects investigate factors influencing the gas mileage a car gets. The variables are: new or old spark plugs, the presence or absence of STP, regular or radial tires, regular or graphite oil, and the brand of gasoline used.

7. Fishing, a variation of the naturalistic task used by Pulos and Linn (1978), allows investigation of factors influencing the number of fish caught. The variables are: the type of bait, location, the presence or absence of a bobber, the type of reel, and the presence or absence of a radio.

8. In Pounds Away, subjects investigate factors influencing weight loss. The variables are: the type of food eaten, the type of exercise, the amount of sleep, the amount of water consumed per day, and the number of meals eaten per day.

9. Soaps, allows investigation of the whiteness of clothes coming out of a washing machine. Variables are: brand of soap, presence or absence of bleach, the water temperature, the presence or absence of fabric softener, and the type of washing machine used.

10. Runaround is a naturalistic task about running speed. Variables are: running location, running surface, clothing worn, amount of sleep, diet, time of day when running, and type of shoes worn.

Questions

The five questions in the controlling variables tasks are described in this section. The constructing, criticizing, planning, and analyzing questions measured reasoning about the controlling variables strategy. Two of these questions, constructing and criticizing, were combined to form a controlling score for data analysis. The fifth question, naming, measured the subjects' expected variables.

1. Constructing measured ability to construct a controlled experiment for a particular variable. Subjects also explained their responses. For example, in Bending Rods, subjects were told: "Do an experiment to show that thickness of the rod makes a difference in how far down the rod bends," and were asked "Why did you do it that way?" Subjects passed if all variables except the variable under investigation were kept the same in two or more trials. Failures were categorized into two groups based on subjects' explanations. These were: (a) subject declared the variable was not important, and when probed, indicated that the uncontrolled variable did not need to be controlled or (b) subjects gave any other reason such as pointing out that they had controlled some variables or had investigated the correct variable.

2. Criticizing measured ability to criticize the procedure for a hypothetical uncontrolled experiment. Subjects passed if they identified the confound.

3. Planning measured how subjects would investigate each of the experimenter's comprehensive variables. For example, in Springs the interviewer asked, "What experiments would you do to find out whether each of these things (pointing to list of experimenter's comprehensive variables) actually makes a difference in how far the Springs expand?" Subjects passed if they conducted one or more controlled experiments.

4. Analyzing measured ability to criticize an experiment when shown the result but not told the procedure. For example, in Bending Rods, two rods painted different colors and made of different materials were placed at unequal lengths protruding from a wooden stand. Hanging equal weights on the rods, the experimenter demonstrated that the short rod bent more than the long rod and asked, "Is this a good experiment to prove that short rods bend more than long rods? Why?" and "Do you have any questions to ask about the experiment?" The length variable is called the alleged causal variable in this problem because the

experiment alleged that length caused the outcome. Subjects were not told that the rods differed in material. Subjects passed if they indicated that some variable might be uncontrolled. Correct responses were categorized either (a) subject stated a controlling variables rule, such as "Everything else has to be the same," or (b) subject pointed out a possible confound, e.g., "I don't know if the rods are made of the same material."

One special analyzing question, called the controlled analysis question, was administered in the autos task. This question was identical to a regular analyzing question except that several potential explanations for the reported outcome were eliminated. (For example, for an experiment about the effects of type of gasoline on mileage subjects were told that the two cars were driven over the same roads).

5. Naming measured the subjects' expected variables. For example, in Pounds Away, the interviewer asked, "If you had to lose weight, what would you do differently than you do now? What might someone else do? What else?" In Bending Rods the apparatus was presented and the interviewer asked, "Tell me what you think might make a difference in how far down these rods bend...What else?" All variables in laboratory tasks were visable. The score was the number of variables named. For all tasks, after the Naming question, subjects were given a card listing the experimenter's comprehensive variables.

Formats of controlling tasks

Four formats were used, one for each controlling question as shown in Table 1.

1. The open format, used for Spinning Wheels and Runway, allowed all levels of each variable to be combined. That is, for Spinning Wheels a large sphere could be placed in any size hole at any distance from the center of the wheel.

2. The fixed, format, used for Bending Rods, Springs, Auto, and Toothpaste, limited the possible combinations of variables. For example, there were eight

springs and eight autos (with their characteristics described) so some combinations of variables were omitted. Thus, in fixed format it was not possible to investigate the material of the spring using short springs with thin wire because both materials were not available for short springs with thin wire. Fixed format was thought to be more difficult than open format because experiments must be designed to fit the available choices.

3. The Record sheet format, used for Soaps, Pounds Away, and Fishing, allowed each condition of all variables to be combined, similar to the open format. Naturalistic tasks did not have objects to manipulate but the record sheet represented the objects pictorially. (e.g. rather than actual fishing rods, there were pictures of fishing rods.) Subjects indicated how they would design each experimental trial on the sheet. Thus, for Soaps, subjects indicated how they would wash two loads of equally soiled clothes to determine whether bleach makes a difference. They indicated which soap they would use for the first load and which they would use for the second load, etc.

4. A unique format, similar to the open format, was used for Runaround. Rather than a record sheet, a chart with the alternative conditions of each variable was used; experimental trials were indicated with red and yellow chips. The subject designed two runs (on different days) to find out if one variable, e.g., sleep, influenced running time. The subject chose a condition for each variable for each trial. The successful subject changed only one variable at a time.

Group Measure to Elicit Subjects' Expected Variables

A paper-and-pencil group-administered survey, "What is Your Opinion?" assessed expectations about each variable in each task. Two questions were asked: (1) Which condition of the variable do you think will have the greatest effect on the outcome, e.g., will the heavy or light weight make the spring expand the most? and (2) How much difference do you think the variable will make?: a big difference, a little difference, no difference, or don't know.

## Results and Discussion

Responses were analyzed to reveal content effects on reasoning. Contents employed were laboratory and naturalistic. Questions guiding the analysis were:

1. What is the interaction between measures of expected variables and laboratory versus naturalistic content?
2. What is the interaction between each of the four controlling variables questions and laboratory versus naturalistic content?
3. What is the interaction between the three task formats and laboratory versus naturalistic content?
4. Are age, sex, or school related to performance?

### Interactions of the subjects' expected variables and laboratory versus naturalistic content

The What is Your Opinion? survey measured both the accuracy and the confidence of the subjects' expected variables. Accuracy of expectations was measured by asking subjects how each of the experimenter's comprehensive variables would effect the outcome, e.g., do heavy or light weights make rods bend more? For laboratory tasks 77% of the answers were correct and for naturalistic tasks 81% of the answers were correct, indicating no significant differences in accuracy of expectations. Confidence of expectations was measured by asking whether each of the experimenter's comprehensive variables would make a "big difference" in the outcome. Subjects expected laboratory variables to make a "big difference" (65% of the time) significantly more often than for naturalistic variables (47% of the time): ($t = 6.21$, $p < .01$). Thus, subjects had stronger expectations about variables for laboratory than for naturalistic tasks.

The subjects' expected variables for each task were measured in the interview by the naming question. Subjects had significantly more expected variables for laboratory than for naturalistic tasks (Table 1), consistent with the stronger expectations mentioned above. The number of subjects' expected variables on one task

did not correlate with the number on other tasks to a high degree as reflected in the low reliability for naming.

Questions about the controlling variables strategy and laboratory versus naturalistic content

Interactions of the four controlling variables questions, constructing and criticizing, were combined because they correlated highly (.82), replicating previous studies (e.g., Linn & Rice, 1979; Linn, Pulos, & Gans, in press). Together they form what we call the controlling score.

Content effects for controlling, planning, and analyzing questions were assessed using average scores for each laboratory and naturalistic task. Means, standard deviations, and reliabilities are given in Table 2 for each content. Reliabilities are generally high; slightly higher standard deviations for naturalistic controlling tasks reflect a slight ceiling effect for laboratory controlling.

Recall that this study included two separate interviews for each subject. Analysis was done separately for each interview and for both interviews combined. No interview effects were found.

Content effects were analyzed using a repeated measures analysis of variance. Results (Table 3) are discussed for controlling, planning, and analyzing questions below. Since male/female differences were never significant, this variable was dropped from further analyses.

---

Insert Tables 2 and 3 about here

---

Content effects on Controlling questions. Recall that controlling includes construction questions and criticizing questions. Subjects must design or criticize experiments.

We found that content influenced controlling. Laboratory tasks were easier than naturalistic tasks as shown in Table 2. Success was high (65 to 80% correct responses). Using analysis of variance (Table 3) we found that more than 12% of the variance in controlling was accounted for by content.

Controlling performance reflects differences in the subjects' expected variables: they control more laboratory than naturalistic variables and also expect more laboratory than naturalistic variables to be important.

Explanations of errors on controlling varied for the two contents. Unsuccessful subjects were asked why a specific variable in their experiment was not controlled. If this variable was not one of their expected variables we anticipated the explanation would be "I don't think that variable influences the outcome". Consistent with the smaller number of expected variables for naturalistic content, these responses were twice as frequent for naturalistic (10%) than for laboratory tasks (5%), a significant difference ($t = 2.93$, $p < .01$). Thus content effects for controlling reflect differences in subjects' expected variables: subjects controlled more laboratory than naturalistic variables. When they failed naturalistic tasks they were likely to claim that the variable omitted did not influence the outcome, presumably because it was not one of their expected variables.

Content effects on planning questions. Recall that planning questions request the subject to plan a series of experiments. As expected, planning is more difficult than controlling: 26% success for naturalistic, 53% for laboratory.

Content effects for planning replicated findings for controlling. In repeated measures analysis of variance, content accounted for 10% of the variance in planning, with laboratory content easier than naturalistic content (Table 3). As for

controlling, content effects reflect differences in the subjects' expected variables: subjects control more laboratory than naturalistic variables.

Content effects on analyzing questions. As described above, analyzing questions require subjects to criticize experiments where they are told an alleged causal variable but are not told whether other variables were controlled. Analysis is more difficult than controlling but easier than planning (52% success on analyzing compared to 72% success on controlling and 41% success on planning).

Content effects for analyzing were found: naturalistic tasks were slightly easier than laboratory tasks (Table 3). Content effects for analyzing questions were consistent with differences in the subjects' expected variables. Successful responders could question the alleged causal relationship, saying: "That experiment is bad because some other variable probably caused the outcome". Subjects had to expect that some other variable caused the outcome in order to question the causal relationship. Subjects who questioned the alleged causal relationship probably did so because the variable alleged to cause the outcome was not one of their expected variables. Naturalistic tasks had fewer expected variables than laboratory tasks. In fact, alleged variables in analyzing questions were subjects expected variables, 92% of the time for laboratory compared to 49% of the time for naturalistic analyzing questions ($t = 8168$, $p < .01$). Therefore, questioning the alleged causal relationship should have been more common for naturalistic than for laboratory tasks, because the alleged variable was less likely to be one of the subject's expected variables.

We verified the impact of the subjects' expected variables on performance by examining reasons for correct responses. Two types of correct responses were possible. One response was to question the alleged causal variable (e.g. "It may not be the length of the rod like you said, but the material that caused the effect").

Another response was to cite a rule about controlling, saying: "I don't know if all the other possible influences were the same". Questioning the alleged causal variable was almost twice as likely for naturalistic analyzing questions (43%) as for laboratory (26%) analyzing questions. This difference was significant ($t = 4.3$, $p < .01$). In contrast, for both contents, citing the rule "everything else should be the same" was equally likely and infrequent (5% of responses). Thus subjects criticize analyzing question results more frequently if the alleged causal variable was not one of their expected variables.

What happened when the subjects' expected variables were not available to explain the analysis question outcome? On the controlled analysis question, subjects were told that expected variables could not have caused the outcome reported in the question. Thus, likely answers to the question were eliminated. On the controlled analyzing question, subjects were dramatically less successful (19% success) than on the regular question (40% success). Few successful responders suggested that another variable caused the outcome since most alternative causes were eliminated by the questions. Successful responders, instead, stated the controlling variables rule. The rule was more likely to be used for this specific question (14%) than for any other analyzing, constructing or criticizing question. It appears that the rule is available but is not usually the first response generated.

Relationship Between Reasoning Questions

Scores for laboratory and naturalistic versions of each question were reasonably correlated: Controlling = .75; Planning = .73; Analyzing = .56. The relationship among controlling, planning, and analyzing was generally high (average correlation = .65). These correlations suggest a common factor in all the reasoning questions.

Interactions between task format and laboratory versus naturalistic content

We hypothesized that fixed format would be the most difficult, and the results bear this out (Tables 4 & 5). The fixed format requires keeping more information in mind than does the open or record format: variables cannot be controlled sequentially. In Bending Rods, for exapmle, the subject might have chosen two brass rods to investigate cross-section only to discover that cross section is confounded with thickness for brass rods. In open or record format subjects can choose conditions for each variable and always find unconfounded tests.

---

Insert Tables 4 and 5 about here

---

Fixed format is more difficult; does format interact with content? It appears not. Content effects performance on fixed format about the same way it effects performance on open and record formats.

Effect of sex, age, and school on performance

As noted above, there were no significant effects for sex so this variable was dropped from further analysis.

The age related improvements in performance on all reasoning questions reflected a developmental trend. Age accounted for the most variance in planning and the least in the controlling, suggesting that planning was more influenced by a developmental mechanism. We investigated three explanations of these findings: (1) across the board acquisition of a formal reasoning strategy; (2) changes in processing capacity as described by de Ribaupierre and Pascual-Leone (1979); or (3) increases in content knowledge.

The first explanation, acquisition of formal reasoning could account for improvement in controlling, planning and analyzing with age if some subjects at each age acquire formal reasoning. How can acquisition of formal reasoning be established? We could consider a subject formal if they correctly used the controlling

strategy at least once but then all subjects would be classified as formal. We could consider a subject at formal if they responded to the planning question correctly since planning resembles the controlling task used by Inhelder and Piaget (1958). The planning question required a more complex strategy, keeping track of several experiments; there were subjects who failed all planning questions, suggesting that planning may develop in this age range. Thus some support for an increase in formal reasoning is possible if the definition of formal is chosen correctly.

Processing capacity explains age related changes in performance by an increase in the amount of information that can be processed. The increasing difficulty of controlling, analyzing, and planning questions reflects the increasing need for processing capacity. As discussed by Linn & Levine (1978), analyzing requires more processing capacity than controlling because the results of the analyzing question must be kept in mind. Planning, since it requires designing a series of experiments also requires more processing capacity than controlling. Thus, an increase in processing capacity may explain the age related increases in reasoning performance as well as increases in the difficulty of the reasoning questions.

Age related increases in content knowledge are reflected in increases in Naming and therefore in the subjects expected variables. Thus, older subjects might perform better because they have more accurate expected variables than younger subjects. Further evidence for content/knowledge effects comes from the relationship between school and reasoning performance: Students in high SES schools outperformed those in low SES schools (Tables 4, 5). School accounted for about 10% of the variance in each reasoning question. Increases in content knowledge are associated with better performance but the causal role of content is uncertain.

Acquisition of formal reasoning and processing capacity could each explain age effects; their relative importance is difficult to ascertain. Content/

knowledge, however, clearly plays a role. Content could have a direct effect or content could indirectly influence performance by influencing strategy application. Another possible indirect content influence could be a reduction of the processing capacity requirements of a task accompanying increases in content knowledge.

In summary, we found that content influenced reasoning. The subject's expected variables influenced performance on controlling variables questions. Subjects expect to consider more laboratory task variables than naturalistic task variables.

Performance on controlling and planning controlling variables questions was more successful for laboratory than for naturalistic tasks. Performance on analyzing controlling variables questions was more successful for naturalistic than for laboratory tasks.

It appears that performance on controlling and planning questions is directly related to the subjects' expected variables. Subjects tend to consider their expected variables in designing experiments, and they have more expected variables for laboratory than for naturalistic tasks. In contrast, it appears that performance on analyzing questions is inversely related to the subjects' expected variables. If the alleged causal relationship in the analyzing question is not about one of the subjects' expected variables then the subject (correctly) queries the causal relationship. Since the alleged causal relationship for laboratory tasks is more likely to be about a subjects' expected variable than it is for naturalistic tasks, analyzing questions are more difficult for laboratory than for naturalistic tasks.

## Implications

These results reveal the strong influence of content on reasoning performance: between 8% and 15% of variance on the controlling variables reasoning questions was associated with content. Contrary to some of Piaget's writings, content effects were systematic. In particular, the subjects' expected variables did influence how the problem was solved.

At the onset we asked: Is it formal if it's not physics? We found that by choosing physics problems and questions which deemphasized content Piaget avoided many of the content influences. Piaget chose laboratory tasks and since most laboratory variables are part of the subjects' expected variables content affects accounted for little variance. Piaget chose questions to assess the controlling variables strategy which emphasized the procedure for a controlled experiment and minimized content. In contrast, our analyzing questions required reasoning about the procedure for a controlled experiment but also required reasoning about the experimental outcome. In analyzing questions when the experimental outcome was emphasized we found that subjects frequently considered the content in their response rather than applying the controlling variables strategy.

Content effects explain why certain reasoning problems are failed by subjects who know the controlling variables strategy: subjects only consider their expected variables in applying the strategy. Thus, subjects apply the strategy correctly to their expected variables but do not consider other variables. In explaining why they did not control a variable, subjects frequently said, "That variable does not make any difference," suggesting that the variable was not one of their expected variables.

Content effects clarify research findings of pervasive inaccurate physics reasoning. McDermott (Note 2) reports inaccurate reasoning about acceleration; Clement (1979) reports inaccurate reasoning about force; Champagne, Klopfer, and Anderson (1979) report inaccurate reasoning about mechanics. If subjects have

inaccurate expected variables they might reason inaccurately because they only reason about their expected variables. In the research reported here, subjects considered only their expected variables thereby omitting important variables. In other situations by reasoning about their expected variables, subjects might include unimportant variables such as including weight as a variable in the oscillation of the pendulum. Thus, studies showing inaccurate physics reasoning may reflect inaccurate expected variables.

Task Characteristics

Our laboratory and naturalistic tasks differ in important respects. These differences are largely ecologically valid. Our laboratory tasks involve apparatus, represent closed systems, and involve some variables, such as weight and length on bending rods, whose effects are well understood by subjects. In contrast, our naturalistic tasks involve printed and verbal information, represent open systems, and involve some variables whose effects are poorly understood by subjects like the effect of amount of water consumed and type of diet on weight loss.

Laboratory tasks such as bending rods are usually encountered in science classes accompanied by apparatus. Usually the task variables can be manipulated by using the apparatus. In contrast, naturalistic tasks such as determining how best to lose weight are encountered in a wide range of situations and are not necessarily accompanied by apparatus. Usually the task variables can be manipulated but may be difficult to systematically investigate because they covary with other variables. For example, weight loss may occur when the subject is feeling optimistic and eating less.

By closed versus open system we mean specified versus unspecified limitations on the task variables. Laboratory tasks are often closed systems where only certain variables can be investigated and where the choices about how to conduct the investigation may be limited (rods may only have two possible thicknesses). In contrast, naturalistic tasks are usually open systems where the possible

variables and choices for those variables seem fairly unlimited.

We have shown that subjects' expected variables differ for laboratory and naturalistic tasks. These tasks differ in important respects. Subsequent research could clarify the precise role of apparatus, openness of the system, and understanding of the variables in laboratory and naturalistic task differences.

Question Characteristics

By adding the analyzing questions to those traditionally asked about the controlling variables strategy we found interactions between content and question. Controlling was easier for laboratory tasks while analyzing was easier for naturalistic tasks. Analyzing is an important question because it more closely mirrors what happens in naturalistic problem solving than does controlling or planning. Reasoners rarely set up controlled experiments to test their ideas because they don't have the time or resources. Instead they are likely to analyze incomplete information. Even in the controlled analysis question subjects only mention the controlling variables strategy when they cannot generate an alternative variable that could explain the outcome. Thus, performance on analyzing may more closely resemble naturally occurring problem solving than does performance on controlling.

Expectations

Our hypothesis that expectations influence reasoning performance is supported by this research. We found that the completeness of the subjects' expected variables influenced performance. How do these expectations develop and how can they be modified?

Subjects develop expectations from experiences with the variables. They successively refine their expectations as new information is encountered. Expectations become situation specific only when forced by pervasive evidence. For example, in the Pendulum subjects expect weight to influence oscillation rate (although weight does not influence oscillation rate) because weight is influential

in most situations (like the Balance Beam). Research has shown that most subjects expect weight to influence the oscillation of the Pendulum and are not easily altered to be situation specific by evidence from controlled experiments (Linn, 1977).

The refinement of expectations could enhance reasoning. Expectations about laboratory and naturalistic tasks are verifiable but rarely verified. Encouraging subjects to verify their expectations might enhance reasoning. Obviously, it is inefficient to verify all expectations so a procedure to aid recognition of potentially inaccurate or incomplete expectations might foster reasoning.

Expectations systematically affect formal reasoning. Subjects may not consider the experimenter's comprehensive variables if they are not among their expected variables. Expectations are difficult to modify since they are rarely verified and are refined by many experiences. One counter experience does not change an expectation. Improvement in formal reasoning may occur when subjects verify their expectations.

Philosopher of science Thomas Kuhn depicts two types of scientific research: normal science and revolutionary science. During normal science, research is guided by a paradigm or theory which describes the expectations of the researchers. Normal scientific research follows the paradigm; researchers gather information to augment the paradigm not to change it. In contrast, during the revolutional science, there is a paradigm shift. Revolutionary science occurs when the previous assumptions, ideas, and even problems are given up and a new paradigm emerges. During these times, expectations are not universally agreed upon. Revolutions involve questioning of expectations; there is controversy in the field about the importance of each variable.

If our tasks can be thought of as exemplars of normal science and of revolutionary science, our laboratory tasks resemble normal science while our naturalistic tasks resemble revolutionary science. For our laboratory tasks, expectations

are strongly held and there is consensus about the procedures for gathering additional information. In contrast, for our naturalistic tasks, there is variability in the expectations about the variables, few variables are expected to influence the outcome, and procedures for gathering new information are controversial. Thus our naturalistic tasks may more closely resemble revolutionary scientific investigations than our laboratory tasks. Naturalistic tasks simulate creative scientific research because there is controversy about expectations for the variables.

Classroom instruction might employ naturalistic rather than laboratory tasks to simulate creative scientific research and to motivate students to go into science.

Kuhn points out that revolutions do not necessarily lead to superior logic on the part of the scientists who participate. Of course, revolutions or paradigm shifts do result in some sorts of intellectual gains. These gains may be in the content or knowledge of the area not in the logic. Siegel (in press) points out that revolutions do not imply logical development as Piaget defined it; increases in content knowledge do not imply increases in formal reasoning ability, for example. However, revolutions do imply changes in expectations about the variables in a particular problem. During revolutions, researchers evaluate their expectations and change some of them. This progression of knowledge may occur in addition to the progression in logic described by Piaget. If reasoners can be taught to go through normal and revolutionary periods then instruction which encourages revolutions may be helpful. Adolescents have many inaccurate expectations which, if investigated, might be altered. Thus, instruction to encourage revolutions would encourage verification of expectations.

Naturalistic tasks are important to insure that science education fosters scientific literacy. Instruction using naturalistic situations would enhance the relevance of the instructional program. It is essential to provide instruction

that helps students evaluate expectations about naturally occurring problems. Laboratory tasks may not require such evaluations. Since it is apparent that evaluation of expeccations is very important for naturalistic problem solving, efforts to choose problems for science instruction to illustrate the role of expectations would enhance scientific literacy.

These results suggest that educational implications generated from Piagetian theory may be incomplete. Content effects require more emphasis. Consistent with Duckworth's (1979) suggestions for younger children, variability in performance on reasoning problems may not be due so much to lack of developmentally based strategies as to lack of instruction in how to combine expectations and strategies.

## Reference Notes


1. Linn, M. C., & Swiney, J. Representation as a factor in formal thought. Adolescent Reasoning Project (ARP) Report No. 22. Lawrence Hall of Science, University of California, Berkeley, 1980.
2. McDermott, L. Investigation of the conceptual understanding in the study of motion among introductory physics students. Paper presented at SESAME (Group in Science & Mathematics Education) seminar, University of California, Berkeley, Fall, 1979.
3. Pulos, S., & Linn, M. C. Modifying subjects' expected variables: Educated Guess. Adolescent Reasoning Project (ARP) Report No. 29. Lawrence Hall of Science, University of California, Berkeley, 1980.

## References


Bransford, J. D., & McCarrell, N. S. A sketch of a cognitive approach to comprehension: Some thoughts about understanding what it means to comprehend. In P. Johnson-Laird & P. C. Wason (Eds.), Thinking. Cambridge: Cambridge University Press, 1977.

Case, R. Structures and strictures: Some functional limitations on the course of cognitive growth. Cognitive Psychology, 1974, 6, 544-573.

Champagne, A. B., Klopfer, L. E., & Anderson, J. H. Factors influencing the learning of classical mechanics. University of Pittsburgh: Learning Research and Development Center, 1979.

Clement, J. Mapping a student's causal conceptions from a problem solving protocol. In J. Lochhead, & J. Clement (Eds.), Cognitive Process Instruction. Philadelphia: Franklin Institute Press, 1979.

Duckworth, E. Either we're too early and they can't learn it or we're too late and they know it already: The dilemma of "applying Piaget". Harvard Educational Review, 1979, 49(3), 297-312.

Inhelder, B., & Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958.

Kuhn, T. S. The Structure of Scientific Revolutions. Chicago: The University of Chicago Press, 1962.

Linn, M. C. Scientific reasoning: Influences on task performance and response categorization. Science Education, 1977, 61, 357-363.

Linn, M. C. Cognitive style, training, and formal thought. Child Development, 1978, 49, 874-877.

Linn, M. C., & Levine, D. I. Adolescent reasoning: Influences of question format and type of variables on ability to control variables. Science Education, 1978, 62(3), 377-388.

Linn, M. C., Pulos S., & Gans, A. Correlates of formal reasoning: Content and problem effects. Journal of Research in Science Teaching, in press.

Linn, M. C., & Rice, M. A measure of scientific reasoning: The springs task. Journal of Educational Measurement, 1979, 16, 55-58.

Pascual-Leone, J., Ammon, P., Goodman, D., & Subelman, I. Piagetian theory and Neo-Piagetian analysis as psychological guides in education. In J. M. Gallagher & J. Easley (Eds.), Knowledge and Development (Vol. 2) Piaget and Education. Plenum Publishing Company, 1978.

Piaget, J. The theory of stages in cognitive development. In D. R. Green, M. P. Ford, and G. B. Flamer (Eds.), Measurement of Piaget. New York: McGraw Hill, 1971.

Piaget, J. Intellectual evolution from adolescence to adulthood. Human Development, 1972, 15, 1-12.

Pulos , S., & Linn, M. C. Pitfalls and pendulum. Formal Operator, 1978, 1(2), 1-11.

Siegel, H. On the parallel between cognitive development and the history of science. Philosophy of the Social Sciences, in press.

Table 1

Controlling Variables Tasks

| Tasks | Interview | Format | Number of Variables | Number of Questions | | | |
|---|---|---|---|---|---|---|---|
| | | | | Naming | Planning | Controlling | Analyzing |
| Laboratory | | | | | | | |
| Bending Rods | II | Fixed | 5 | 1 | 0 | 3 | 1 |
| Springs | I | Fixed | 5 | 1 | 1 | 2 | 1 |
| Spinning Wheels | II | Open | 4 | 0 | 0 | 1 | 1 |
| Runway | I | Open | 6 | 0 | 1 | 2 | 0 |
| Naturalistic | | | | | | | |
| Toothpaste | I | Fixed | 5 | 0 | 1 | 2 | 1 |
| Autos | II | Fixed | 5 | 0 | 0 | 1 | 1 |
| Fishing | II | Record Sheet | 5 | 1 | 0 | 1 | 1 |
| Pounds Away | II | Record Sheet | 5 | 1 | 0 | 1 | 1 |
| Soaps | II | Record Sheet | 5 | 0 | 0 | 1 | 1 |
| Runaround | I | Unique | 7 | 0 | 1 | 2 | 0 |

Table 2

Performance On All Questions

| Question | No. of Quest. | ALL | | AGE | | | | | | SCHOOL | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 7 | | 9 | | 11 | | Upper Middle | | Middle | | Lower Middle | |
| | | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD |
| Controlling[1] | | | | | | | | | | | | | | | |
| Laboratory I | 4 | 79 | 23 | 65 | 31 | 83 | 17 | 88 | 15 | 91 | 13 | 76 | 13 | 70 | 26 |
| Laboratory II | 4 | 80 | 25 | 71 | 30 | 79 | 23 | 89 | 16 | 90 | 16 | 80 | 24 | 69 | 28 |
| Naturalistic I | 4 | 65 | 32 | 51 | 34 | 65 | 34 | 78 | 24 | 82 | 26 | 65 | 32 | 46 | 30 |
| Naturalistic II | 4 | 67 | 35 | 58 | 36 | 68 | 35 | 74 | 34 | 87 | 20 | 67 | 37 | 47 | 35 |
| Total | 16 | 72 | 26 | 59 | 29 | 74 | 23 | 83 | 20 | 87 | 14 | 72 | 26 | 57 | 26 |
| Reliability | | 89 | | 91 | | 91 | | 87 | | 86 | | 95 | | 85 | |
| Planning[1] | | | | | | | | | | | | | | | |
| Laboratory I | 2 | 53 | 40 | 34 | 35 | 50 | 39 | 73 | 37 | 67 | 37 | 57 | 41 | 35 | 37 |
| Naturalistic I | 2 | 28 | 39 | 17 | 31 | 19 | 34 | 48 | 44 | 43 | 42 | 29 | 40 | 13 | 29 |
| Total | 4 | 41 | 34 | 26 | 23 | 35 | 32 | 61 | 36 | 56 | 34 | 44 | 36 | 24 | 25 |
| Reliability | | 89 | | 72 | | 90 | | 92 | | 90 | | 91 | | 77 | |
| Analyzing[1] | | | | | | | | | | | | | | | |
| Laboratory II | 2 | 42 | 42 | 26 | 32 | 41 | 42 | 57 | 45 | 68 | 39 | 32 | 39 | 25 | 35 |
| Naturalistic II | 4 | 57 | 38 | 44 | 40 | 55 | 40 | 71 | 29 | 68 | 35 | 56 | 38 | 46 | 40 |
| Total | 6 | 52 | 35 | 38 | 35 | 51 | 35 | 67 | 31 | 68 | 33 | 48 | 33 | 39 | 35 |
| Reliability | | 82 | | 85 | | 83 | | 72 | | 81 | | 80 | | 85 | |
| Laboratory | 3 | 2.86 | .69 | 2.59 | .64 | 2.74 | .69 | 3.24 | .57 | 3.06 | .57 | 2.81 | .87 | 2.70 | .55 |
| Naturalistic | 2 | 2.20 | .84 | 1.89 | .74 | 2.16 | .76 | 2.55 | .92 | 2.43 | .89 | 2.16 | .73 | 2.02 | .88 |
| Total | 5 | 2.60 | .63 | 2.31 | .52 | 2.51 | .59 | 2.96 | .59 | 2.81 | .59 | 2.55 | .66 | 2.44 | .60 |
| Reliability | | 59 | | 31 | | 59 | | 49 | | 47 | | 62 | | 64 | |

[1]Mean percent success, decimals omitted

[2]Mean number of variables named

Table 3

Comparison of Laboratory and Naturalistic Content Using

Analysis of Variance

| Questions | Effects | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Content | | | Age | | | School | | | Significant Interactions | | |
| | F | Sig. | $\omega^2$ | F | Sig. | $\omega^2$ | F | Sig. | $\omega^2$ | F | Sig. | $\omega^2$ |
| Controlling I | 35.64 | <.01 | .14 | 6.44 | <.01 | .08 | 7.44 | <.01 | .11 | 3.06 | <.05 | .02[1] |
| Controlling II | 24.33 | <.01 | .12 | 2.38 | ns | | 9.24 | <.01 | .18 | 3.80 | <.05 | .02[1] |
| Controlling I/II | 44.48 | <.01 | .20 | 4.45 | <.05 | .02 | 8.71 | <.01 | .15 | 5.07 | <.01 | .03[1] |
| Controlling (Fixed Format Only) | 31.42 | <.01 | .12 | 4.38 | <.05 | .02 | 9.27 | <.01 | .16 | | | |
| Planning | 28.68 | <.01 | .10 | 9.83 | <.01 | .17 | 6.32 | <.01 | .11 | | none | |
| Analyzing | 14.70 | <.01 | .08 | 6.11 | <.01 | .11 | 6.61 | <.01 | .12 | | none | |
| Naming | 46.06 | <.01 | .21 | 7.13 | <.01 | .13 | 2.27 | ns | | | none | |

[1]School x Question Content

Table 4

Performance On Controlling Questions in Each Format

| | ALL | | Age | | | | | | School | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | 7 | | 9 | | 11 | | Upper Middle | | Middle | | Lower Middle | |
| | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD | $\bar{X}$ | SD |
| **Laboratory** | | | | | | | | | | | | | | |
| Fixed Format | 7.68 | 2.34 | 6.59 | 2.91 | 7.72 | 2.05 | 8.68 | 1.47 | 8.82 | 1.39 | 7.71 | 2.44 | 6.50 | 2.49 |
| Open Format | 8.33 | 2.40 | 7.08 | 3.12 | 8.72 | 1.79 | 9.07 | 1.75 | 9.35 | 1.13 | 7.96 | 2.47 | 7.74 | 2.94 |
| **Naturalistic** | | | | | | | | | | | | | | |
| Fixed Format | 6.15 | 3.31 | 5.33 | 3.33 | 5.89 | 3.38 | 7.18 | 3.07 | 7.96 | 2.40 | 5.96 | 3.16 | 4.44 | 3.42 |
| Record Format | 6.73 | 3.71 | 5.67 | 3.63 | 7.10 | 3.70 | 7.36 | 3.73 | 8.79 | 2.20 | 6.75 | 3.86 | 4.64 | 3.72 |

Table 5

Comparison of Formats Within Laboratory and Naturalistic Content

| | Effects | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Format | | | Age | | | School | | | Significant Interactions[1] | | |
| | F | Sig. | $r^2$ | F | Sig. | $r^2$ | F | Sig. | $r^2$ | F | Sig. | $r^2$ |
| **Laboratory Controlling Questions** | | | | | | | | | | | | |
| Fixed versus Open format | 11.47 | < .01 | | 5.46 | <.01 | .10 | 5.86 | <.01 | .11 | | none | |
| **Naturalistic Controlling Questions** | | | | | | | | | | | | |
| Fixed versus Record format | 3.96 | <.05 | | 1.85 | ns | | 10.70 | <.01 | .19 | | none | |

[1] only those accounting for one or more percent of the variance are noted.